UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/520,678	03/07/2000	Michael Tipping	1018.093US1	9074

7590 09/05/2003

Law Offices of Albert S. Michalik
704-228th Avenue NE
Suite 193
Sammamish, WA 98074

EXAMINER	
DAY, HERNG DER	
ART UNIT	PAPER NUMBER
2123	10

DATE MAILED: 09/05/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

4

Office Action Summary	Application No.	Applicant(s)
	09/520,678	TIPPING ET AL.
	Examiner	Art Unit
	Herng-der Day	2123

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *14 June 2003* .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-19* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-19* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *07 March 2000* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

1. This communication is in response to Applicants' Amendment (paper # 8) to Office Action dated January 15, 2003 (paper # 6), mailed June 14, 2003.

1-1. Claims 1, 3, 6, 10, and 12-16 have been amended; claims 1-19 are pending.

1-2. Claims 1-19 have been examined and claims 1-19 have been rejected.

Drawings

2. The drawings are objected to for the following reasons. A proposed drawing correction or corrected drawings are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

2-1. Figures 1 and 2 should be designated by a legend such as --Prior Art-- because only that which is old is illustrated. See MPEP § 608.02(g).

Claim Rejections - 35 USC § 101

3. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4. Claims 1, 10, and 13 are rejected under 35 U.S.C. 101 because the inventions as disclosed in claims are directed to non-statutory subject matter.

4-1. Claims 1, 10, and 13 recite a series of steps describe mathematical algorithms of manipulating various distributions to approximate a posterior distribution for modeling a data set. These steps neither recite data gathering limitations or post-mathematical operations that

might independently limit the claims beyond the performance of a mathematical operation nor limit the use of any output to a practical application providing a useful, concrete, and tangible result. In other words, none of the claims possess limitations that eliminate the rejection for lack of statutory subject matter. Hence, the claims are unpatentable. See In *re Warmerdam,* 33 F.3d 1354, 1360 (Fed. Cir 1994).

4-2. The Examiner acknowledges that even though the claims are presently considered non-statutory they are additionally rejected below over the prior art. The Examiner assumes the Applicants will amend the claims to overcome the 101 rejections and thus make the claims statutory.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Matsumoto et al., "Reconstruction and Prediction of Nonlinear Dynamical System: A Hierarchical Bayes Approach with Neural Nets", ICASSP '99, Proceedings of 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, March 1999, Volume 2, pages 1061-1064, in view of Simoudis et al., U.S. Patent 5,692,107 issued November 25, 1997.

6-1. Regarding claims 1-5, Matsumoto et al. disclose a computer-implemented method for modeling a data set comprising:

(Claim 1) selecting an initial set of hyperparameters for determining a prior distribution for the data set for modeling thereof, the prior distribution approximated by a product of a distribution of the set of hyperparameters, a distribution of a set of weights, and a distribution of a set of predetermined additional parameters (equation (1.1), column 1, page 1061); and,

such that the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of predetermined additional parameters as have been iteratively updated approximates the posterior distribution for modeling of the data set for probabilistic prediction (equation (2.8), column 2, page 1062);

(Claim 4) the data set comprises a continuous data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for noise (ν is a noise process, column 1, line 5, page 1062);

(Claim 5) the data set comprises a discrete data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for a lower bound (equation (1.2), column 1, page 1061).

Matsumoto et al. fail to expressly disclose: (1) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached; (2) initially inputting the data set to be modeled; and (3) outputting the posterior distribution.

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques in order to extract a predictive model (Simoudis, column 3, lines 15-25). Depending on the specific application, the predictive model

may include, for example, a trained neural network for neural models, and the like (Simoudis, column 4, lines 53-56). The user sets module-specific parameters. One of ordinary skill in the art of extracting predictive models and inductive learning knows that the module-specific parameters may include the criteria of convergence, which is a design choice. Once the user determined that the results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (Simoudis, column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(Claim 1) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached (selects 203, generated 204, set 205, applied 206, and examination 207, column 4, lines 43-53);

(Claim 2) initially inputting the data set to be modeled (a data source 114 is selected 200 and input into the system, column 4, lines 36-37);

(Claim 3) outputting the posterior distribution as approximated by the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of predetermined additional parameters (the extracted predictive model then may be saved 209).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Matsumoto et al. to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 1-5 because, as Simoudis et al. suggest, the modules may be custom designed for specific applications (Simoudis, column 3, lines 26-27),

i.e., a application of approximating a posterior distribution of equation (2.8) for probabilistic prediction.

6-2. Regarding claims 6-9, Matsumoto et al. disclose a computer-implemented method for modeling a data set comprising:

(Claim 6) obtain a posterior distribution for the data set (equation (2.8), column 2, page 1062);

(Claim 7) selecting an initial set of hyperparameters for determining the prior distribution for the data set, the prior distribution approximated by a product of a distribution of the set of hyperparameters, a distribution of a set of weights, and a distribution of a set of predetermined additional parameters (equation (1.1), column 1, page 1061); and,

(Claim 8) the data set comprises a discrete data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for a lower bound (equation (1.2), column 1, page 1061);

(Claim 9) the data set comprises a continuous data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for noise (ν is a noise process, column 1, line 5, page 1062).

Matsumoto et al. fail to expressly disclose: (1) inputting a data set to be modeled; (2) determining a relevance vector learning machine via a variational approach; (3) outputting at least the posterior distribution; and (4) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached.

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques in order to extract a predictive model (Simoudis, column 3, lines 15-25). Depending on the specific application, the predictive model may include, for example, a trained neural network for neural models, and the like (Simoudis, column 4, lines 53-56). The user sets module-specific parameters. One of ordinary skill in the art of extracting predictive models and inductive learning knows that the module-specific parameters may include the criteria of convergence, which is a design choice. Once the user determined that the results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (Simoudis, column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(Claim 6) inputting a data set to be modeled (Simoudis, a data source 114 is selected 200 and input into the system, column 4, lines 36-37);

determining a relevance vector learning machine via a variational approach (Simoudis, inductive module 104', Fig. 1); and,

outputting at least the posterior distribution for the data set for probabilistic prediction (Simoudis, the extracted predictive model then may be saved 209);

(Claim 7) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached (Simoudis, selects 203, generated 204, set 205, applied 206, and examination 207, column 4, lines 43-53).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Matsumoto et al. to incorporate the teachings of Simoudis

et al. to obtain the invention as specified in claims 6-9 because, as Simoudis et al. suggest, the modules may be custom designed for specific applications (Simoudis, column 3, lines 26-27), i.e., a application of approximating a posterior distribution of equation (2.8) for probabilistic prediction.

6-3. Regarding claims 10-12, Matsumoto et al. disclose a machine-readable medium having instructions stored thereon for execution by a processor to perform a method for modeling a continuous data set comprising:

(Claim 10) selecting an initial set of hyperparameters for determining a prior distribution for the continuous data set for modeling thereof, the prior distribution approximated by a product of a distribution of the set of hyperparameters, a distribution of a set of weights, and a distribution of a set of parameters accounting for noise (equation (1.1), column 1, page 1061; ν is a noise process, column 1, line 5, page 1062); and,

such that the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of parameters accounting for noise as have been iteratively updated approximates the posterior distribution for modeling of the continuous data set for probabilistic prediction (equation (2.8), column 2, page 1062);

Matsumoto et al. fail to expressly disclose: (1) interactively updating the distributions; (2) initially inputting the data set to be modeled; and (3) outputting the posterior distribution.

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques in order to extract a predictive model (Simoudis, column 3, lines 15-25). Depending on the specific application, the predictive model may include, for example, a trained neural network for neural models, and the like (Simoudis,

column 4, lines 53-56). The user sets module-specific parameters. One of ordinary skill in the art of extracting predictive models and inductive learning knows that the module-specific parameters may include the criteria of convergence, which is a design choice. Once the user determined that the results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (Simoudis, column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(Claim 10) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of parameters accounting for noise until a predetermined convergence criterion has been reached (selects 203, generated 204, set 205, applied 206, and examination 207, column 4, lines 43-53);

(Claim 11) initially inputting the continuous data set to be modeled (a data source 114 is selected 200 and input into the system, column 4, lines 36-37);

(Claim 12) outputting the posterior distribution as approximated by the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of parameters accounting for noise (the extracted predictive model then may be saved 209).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Matsumoto et al. to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 10-12 because, as Simoudis et al. suggest, the modules may be custom designed for specific applications (Simoudis, column 3, lines 26-27), i.e., a application of approximating a posterior distribution of equation (2.8) for probabilistic prediction.

6-4. Regarding claims 13-15, Matsumoto et al. disclose a machine-readable medium having instructions stored thereon for execution by a processor to perform a method for modeling a discrete data set comprising:

(Claim 13) selecting an initial set of hyperparameters for determining a prior distribution for the discrete data set for modeling thereof, the prior distribution approximated by a product of a distribution of the set of hyperparameters, a distribution of a set of weights, and a distribution of a set of parameters accounting for a lower bound (equation (1.1), column 1, page 1061; equation (1.2), column 1, page 1061); and,

such that the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of parameters accounting for a lower bound as have been iteratively updated approximates the posterior distribution for modeling of the discrete data set for probabilistic prediction (equation (2.8), column 2, page 1062);

Matsumoto et al. fail to expressly disclose: (1) interactively updating the distributions; (2) initially inputting the data set to be modeled; and (3) outputting the posterior distribution.

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques in order to extract a predictive model (Simoudis, column 3, lines 15-25). Depending on the specific application, the predictive model may include, for example, a trained neural network for neural models, and the like (Simoudis, column 4, lines 53-56). The user sets module-specific parameters. One of ordinary skill in the art of extracting predictive models and inductive learning knows that the module-specific parameters may include the criteria of convergence, which is a design choice. Once the user determined that the results are satisfactory based on the user's queries or hypotheses, a predictive

model is extracted based on such results (Simoudis, column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(Claim 13) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of parameters accounting for a lower bound until a predetermined convergence criterion has been reached (selects 203, generated 204, set 205, applied 206, and examination 207, column 4, lines 43-53);

(Claim 14) initially inputting the discrete data set to be modeled (a data source 114 is selected 200 and input into the system, column 4, lines 36-37);

(Claim 15) outputting the posterior distribution as approximated by the product of the distribution of the set of hyperparameters, the distribution of the set of weights, and the distribution of the set of parameters accounting for a lower bound (the extracted predictive model then may be saved 209).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Matsumoto et al. to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 13-15 because, as Simoudis et al. suggest, the modules may be custom designed for specific applications (Simoudis, column 3, lines 26-27), i.e., a application of approximating a posterior distribution of equation (2.8) for probabilistic prediction.

6-5. Regarding claims 16-19, Matsumoto et al. disclose a machine-readable medium having instructions stored thereon for execution by a processor to perform a method comprising:

(Claim 16) obtain a posterior distribution for the data set for probabilistic prediction (equation (2.8), column 2, page 1062);

(Claim 17) selecting an initial set of hyperparameters for determining the prior distribution for the data set, the prior distribution approximated by a product of a distribution of the set of hyperparameters, a distribution of a set of weights, and a distribution of a set of predetermined additional parameters (equation (1.1), column 1, page 1061); and,

(Claim 18) the data set comprises a discrete data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for a lower bound (equation (1.2), column 1, page 1061);

(Claim 19) the data set comprises a continuous data set, such that the set of predetermined additional parameters comprises a set of parameters accounting for noise (ν is a noise process, column 1, line 5, page 1062).

Matsumoto et al. fail to expressly disclose: (1) inputting a data set to be modeled; (2) determining a relevance vector learning machine via a variational approach; (3) outputting at least the posterior distribution for the data set; and (4) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached.

Simoudis et al. teaches the selection of a data analysis module from several different modules implementing different data mining techniques in order to extract a predictive model (Simoudis, column 3, lines 15-25). Depending on the specific application, the predictive model may include, for example, a trained neural network for neural models, and the like (Simoudis, column 4, lines 53-56). The user sets module-specific parameters. One of ordinary skill in the art of extracting predictive models and inductive learning knows that the module-specific

parameters may include the criteria of convergence, which is a design choice. Once the user determined that the results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (Simoudis, column 4, lines 42-57). Specifically, Simoudis et al. disclose the missing elements:

(Claim 16) inputting a data set to be modeled (Simoudis, a data source 114 is selected 200 and input into the system, column 4, lines 36-37);

determining a relevance vector learning machine via a variational approach (Simoudis, inductive module 104', Fig. 1); and,

outputting at least the posterior distribution for the data set (Simoudis, the extracted predictive model then may be saved 209);

(Claim 17) interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached (Simoudis, selects 203, generated 204, set 205, applied 206, and examination 207, column 4, lines 43-53).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teachings of Matsumoto et al. to incorporate the teachings of Simoudis et al. to obtain the invention as specified in claims 16-19 because, as Simoudis et al. suggest, the modules may be custom designed for specific applications (Simoudis, column 3, lines 26-27), i.e., a application of approximating a posterior distribution of equation (2.8) for probabilistic prediction.

Applicant's Arguments

7. Applicants argue the following:

(1) Neither Hearst nor Cortes, whether considered alone or in any permissible combination, discloses all the claimed subject matter (pages 8-14, paper # 8).

(2) Applicants have amended claims 1, 6, 10, 13, and 16 to make explicit what was implicitly recited by these claims to overcome the rejections of claims 1-12 and 14-19 under 35 U.S.C. 101 (pages 14-15, paper # 8).

Response to Arguments

8. Applicants' arguments have been fully considered.

8-1. Response to Applicants' argument (1). Applicants' arguments are persuasive. Therefore, the rejections under 35 U.S.C. 103(a) in paper # 6 have been withdrawn. However, upon further consideration, a new ground(s) of rejection is made, as detailed in sections **6** to **6-5** above.

8-2. Response to Applicants' argument (2). Applicants' arguments are unpersuasive. Claims 1, 10, and 13 are still rejected under 35 U.S.C. 101, as detailed in section **4** above.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Reference to Bishop et al., U.S. Patent 6,556,960 issued April 29, 2003, and filed September 1, 1999, is cited as disclosing a variational inference engine for probabilistic graphical models.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Herng-der Day whose telephone number is (703) 305-5269. The examiner can normally be reached on 9:00 - 17:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Kevin J Teska can be reached on (703) 305-9704. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 746-7239 for regular communications and (703) 746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-3900.

Herng-der Day
August 26, 2003

Hugh Jones

HUGH JONES Ph.D.
PRIMARY PATENT EXAMINER
TECHNOLOGY CENTER